

160-TP-012-001

EP7 Evaluation Plan

**Technical paper - Not intended for formal review
or Government approval.**

Technical Paper

July 1996

Prepared Under Contract NAS5-60000

RESPONSIBLE ENGINEER

Janice Poston Day /s/	7/22/96
Janice M. Poston Day, Science Specialist	Date
EOSDIS Core System Project	

SUBMITTED BY

Joy Colucci /s/	7/24/96
Joy Colucci, Science Office Manager	Date
EOSDIS Core System Project	

Hughes Information Technology Systems
Upper Marlboro, Maryland

This page intentionally left blank.

Abstract

EP7 will be released for evaluation in July 1996. This paper details the evaluation methods that will be employed to gather user feedback on the World Wide Web (WWW) and X/Motif portions of the Evaluation Package.

Usability testing and an on-line user survey, the Comment Survey Tool (CST) have been chosen as the two methods of evaluation. Usability testing will be conducted on both the WWW and X/Motif portions of EP7 at the Landover ECS Development Facility in a controlled environment that allows for observed and measured user response for evaluation of design efficiency. Use of the Comment Survey Tool provides the opportunity to measure important aspects of user needs, preferences, and comments without significant investments in time and workspace. This method of evaluation is designed so that evaluators can access the EP7 at their own convenience.

Evaluators have been selected by ESDIS and DAAC personnel. They have a wide range of science backgrounds and research interests.

The data analysis of the usability testing, the survey, and the comparison between the two will each be summarized in the EP7 Evaluation Results Report, which will be published by the end of September 1996.

Keywords: Evaluation Package 7, EP7, Usability Testing, Survey, CST, Comment Survey Tool

This page intentionally left blank.

Contents

1. Introduction

1.1	Purpose.....	1-1
1.2	Organization.....	1-1

2. Overview

2.1	What is an Evaluation Package?	2-1
2.2	EP7 Evaluation Methods Overview.....	2-1

3. Background

3.1	Key Terms Used in This Document.....	3-1
3.2	Lessons Learned From Previous Evaluations.....	3-1

4. Comment Survey Tool

4.1	Survey Methodology.....	4-1
4.2	Survey Goals.....	4-2
4.3	Survey Participants	4-2
4.4	Survey Schedule.....	4-2
4.5	Survey Questions	4-3

5. Usability Testing

5.1	Test Methodology	5-1
5.2	Test Goals	5-1
5.3	Test Participants.....	5-2
5.4	Test Schedule	5-2

5.5	Test Environment.....	5-2
5.6	Test Tasks	5-4
5.7	Coordination with Comment Survey Tool.....	5-4

6. Demonstration Scripts

6.1	Overview Demonstration for Usability Participants.....	6-1
6.2	Remote Evaluator Demonstration.....	6-1

7. References 1

Figures

5-1.	Usability Test Environment.....	5-3
------	---------------------------------	-----

1. Introduction

1.1 Purpose

This paper describes the two evaluation methods that will be used to evaluate Evaluation Package 7 (EP7): usability testing and user survey. Descriptions for the implementation of usability tests for EP7 and the on-line user survey, known as the Comment Survey Tool (CST), are provided. The data analysis schedule for these two data collection methods is given.

1.2 Organization

This paper is organized as follows:

Chapter 2 introduces the Evaluation Package 7 (EP7), usability testing, and the Comment Survey Tool (CST). Chapter 3 provides definitions of terminology used throughout the document as well as examples of lessons learned from previous usability test sessions. Information on the CST as well as the EP7 results and data distribution policy are provided in Chapter 4. Usability Testing methodology, goals, environment, and tasks are detailed in Chapter 5. Chapter 6 relates to the Demonstration Scripts that will be employed by both Usability Participants and Remote EP7 Evaluators to learn about EP7 and view some of the capabilities of the system.

Questions regarding technical information contained within this Paper should be addressed to the following ECS contacts:

- ECS Contacts
 - Keith Bryant, EP7 Manager, (301) 925-1126, kbryant@eos.hitc.com
 - Janice Poston Day, EP7 Usability Test Lead, (301) 925-0811, jposton@eos.hitc.com

Questions concerning distribution or control of this document should be addressed to:

Data Management Office
The ECS Project Office
Hughes Applied Information Systems
1616 McCormick Dr.
Upper Marlboro, MD 20774-5372

This page intentionally left blank.

2. Overview

2.1 What is an Evaluation Package?

A multi-track development process has been adopted by the ECS Team to build the EOSDIS Core System. This process includes the development of a portion of ECS on an incremental track and parallel development of the remainder of ECS on a formal track using traditional waterfall development methodology. To accelerate and accomodate early user feedback, and integrate early prototypes and incremental prototypes, a delivery mechanism called an Evaluation Package (EP) was devised.

An Evaluation Package is a collection of prototypes that allow ECS developers to test concepts, implementation designs, and develop increasingly mature and robust software for the ECS user interface. Evaluation Packages also allow developers to evaluate the advertised capabilities of Commercial-Off-The-Shelf (COTS) software and hardware in the context of ECS. In addition to designing and constructing the user interface for the EP, the “back end” software and hardware that users do not see (e.g., security and communications software, database design, database population) is also developed. Once complete, the EP is delivered to the DAACs and a set of NASA sponsored evaluators, nick-named “Tirekickers” for evaluation. Each EP builds upon and expands the capabilities of the previous EPs, until the last EP in the series, EP7. EP7 will be migrated to the formal track for integration, acceptance testing, and delivery in ECS Release B.

Evaluation Packages focus on the ECS subsystems with heavy user interaction: Client, Interoperability, Data Management, and Data Server subsystems. The EP as shipped to the DAACs and evaluators represents the prototype ECS Client. Using the Client applications users can search for data, view data, order data, manage their user environment, and access the data stored in the Data Server.

2.2 EP7 Evaluation Methods Overview

Evaluation Package 7 will be released for evaluation in July 1996; the evaluation will continue through the middle of September 1996. EP7 will be released in both a World Wide Web and X/Motif versions. The World Wide Web version of EP7 will be released prior to the X/Motif version, but the X/Motif version will have several functionalities not available in the WWW version, therefore, evaluators will be encouraged to use both versions. Evaluation of EP7 will be begin as soon as software are available. EP7 will be evaluated using two methods: Usability Testing and an on-line user survey known as the Comment Survey Tool (CST). The usability test is administered in a controlled environment that allows for observed and measured user response for evaluation of design efficiency. The second evaluation method allows evaluators who examine the EP in an uncontrolled environment (e.g., their office, a university computer lab) to input their user preferences and suggestions to the CST.

Usability tests are designed to capture very detailed data about a user's experience with the EP; they are time and resource intensive. The usability tests are administered in a controlled environment that allow for observed and measured user response. The usability test consists of a series of tasks that the test participants are asked to complete. After each task, they are asked to fill out a user survey. Data about the participant's previous experience with the evaluation of an EP, previous computing experience, and experience using WWW browsers.

Because usability testing is conducted in a controlled environment, variables such as time to access the system, order of user tasks, the amount of help provided can be controlled, in addition, reliable user statistics can be collected. One of the most notable advantages of usability testing is that it enables the system developers to directly observe the way users use the system. The drawbacks to this method are that it is time consuming and can generate a significant amount of data. Because of these factors, test sessions must be carefully prepared for and smoothly orchestrated for maximum effectiveness.

The on-line user survey provides the opportunity to measure important aspects of user needs, preferences, and comments with minimal investment in time and workspace. Both the Usability Participants and the Remote Evaluators will use the on-line survey. Collection of data via on-line survey allows users to access the EP7 at their own convenience and alleviates the requirements for a dedicated test environment, travel, etc. Another advantage is that the evaluators will be able and encouraged to access EP7 a number of times during the evaluation period. This will provide information on how users' perceptions of EP7 change over time, the amount of time it takes to learn how to operate various EP7 functions, and how use of EP7 at different times of the day and with different system loads affects user perception of EP7. However, the users entering their comments into the CST will each have a unique experience using EP7, therefore it may be difficult to isolate some perceived and actual problems the users encountered using the system.

For a description of EP7 functional content and physical configuration, consult the following:

- EP7 Documents (<http://observer.gsfc.nasa.gov/ep7/docs.html>)

3. Background

3.1 Key Terms Used in This Document

Remote Evaluators: NASA Tirekickers and other potential end users of the ECS. They will access EP7 at their own convenience by remotely logging in to the X/Motif Client installed at the Landover EDF or accessing the JAVA client via the WWW.

Usability Participants: NASA Tirekickers and other potential end users of the ECS who will participate in the usability test sessions conducted at the Landover EDF. Usability test sessions will be conducted in a controlled environment using tasks developed for EP7 user scenarios.

Evaluation Period: the approximately six week period (July 30 to September 15) during which the Evaluators will be able to log on and test the EP7 and during which, usability tests will be conducted. Evaluators will be asked to access EP7 at least twice during this period.

3.2 Lessons Learned From Previous Evaluations

Usability testing and the CST (formerly known as the Interactive Evaluation Tool) have been used for evaluation of EP3, EP4, and EP6. [Note: there was no EP5]. Both methods were effective in providing valuable data to developers. The lessons learned through use of these evaluation methods, as well as the experiences gained during the evaluation of Prototype Workshop 1 (PW1) and Prototype Workshop 2 (PW2) have been incorporated into evaluation plan for EP7. Many of the lessons learned are not relevant to discussion at this point in the document but for illustrative purposes two examples are given.

- During the usability testing of EP6 usability participants were asked to complete a task and then the related questions in the Comment Survey Tool. The alternative was to have the users complete all of the usability tasks and then answer all of the survey questions at the end of the session. Users liked answering the survey questions as they completed each task. The remaining problem with the task-survey process and the survey in general was that it was too long and contained many redundant questions. To improve the survey, the quality of the survey questions, and the Evaluator response rate, the survey will be streamlined so that fewer, more relevant questions are provided for the Evaluator to answer.
- Although usability participants evaluate EPs using a science scenario, Evaluators who remotely log on to evaluate the system do not have access to such a context-setting scenario. During the EP6 Evaluation Period the ECS DAAC liaisons requested that context-setting science scenarios be provided to aid the evaluation by those Evaluators remotely logging in to the system. Context scenarios for the evaluation of EP7 will be provided.

This page intentionally left blank.

4. Comment Survey Tool

4.1 Survey Methodology

The Comment Survey Tool (CST) is an on-line survey that EP7 users will be able access during the course of their evaluation. It contains a series of questions relating to specific functions of EP7, as well as a few general questions on system as a whole. In addition, a field is available at the bottom of each CST window for users to enter their comments in free-text form. Users can use this comment window to expand on their answers to the survey questions and comment on the features and capabilities of EP7 that they liked and disliked. The Comment Survey Tool will be used by both the Remote Evaluators and the Usability Participants.

The CST questionnaire is developed using input from the Science Office, EP7 developers, EP7 management. Where appropriate, survey questions from previous EPs and PWs are included. The draft questionnaire is tested on a subset of the EP7 evaluator population; questions are modified and retested until consensus among those providing survey questions is reached. The questions are developed in close cooperation with those responsible for usability testing. This is to ensure that the questions asked in the CST reflect the experience of the Usability Participants and Remote Evaluators. The questions are inserted into the CST and the data are downloaded from the CST each week.

To protect the integrity of the evaluation, several items relative to attribution of comments must be maintained. These are as follows:

- Each registered evaluator is given a unique user name and password. Comments and survey information are linked to the user who logged in.
- Non-registered evaluators may log-in using the “guest” account. Guest user comments will be evaluated and the survey data included. Guest user comments can not be attributed to a specific individual.
- To protect the integrity of the evaluation, when the comments and survey results are reported, individual’s names will not be linked to specific comments.
- Upon request, a listing of comments made by an evaluator will be provided but only to the evaluator. The evaluator is free to publish his or her comments.
- During the analysis of the survey data, groupings of comments or survey analysis may be used to determine group unique tendencies in the data. Examples of groups might be a specific DAAC or users services evaluators as a group.
- Summary of the results, indications for changes, and plans for implementing the changes will be presented in an EP7 Results Report.

4.2 Survey Goals

The goal of the on-line survey is to measure users' perception of EP7 system, these include user comments that help to identify user needs and preferences. The user perception will give the development team ideas for future ECS design and functionality.

4.3 Survey Participants

Most of those evaluating EP7 will be science users, primarily earth scientists with a wide range of computer and research experience. Not only do the Evaluators have diverse backgrounds, but they are also dispersed throughout the United States.

There are approximately 100 EP7 Evaluators. They have been selected by DAAC managers, Version 0 representatives, and NASA personnel. Evaluators fall into at least one of the following categories:

- Tirekickers
- Science User Working Group members
- DAAC User Services Working group members
- V0 Development Team members
- Instrument Teams
- Interdisciplinary Investigation Teams
- Graduate Students in the Earth Sciences
- ECS Science Liaisons
- ECS System Engineering Liaisons
- DAAC M&O
- ECS M&O

The evaluators' skills provide the EP7 team with additional expertise to refine the ECS design so that it meets customers' needs. Of these, approximately 15 to 20 will be Usability Participants who will evaluate EP7 at the Landover EDF. The usability testing portion of EP7 Evaluation is discussed in the next chapter.

4.4 Survey Schedule

Each week the EP7 team will export the CST survey data from a centralized database. There will be a once per week data collection process and data integrity check on the user survey data.

To download the information an EP7 team member will execute a procedure against the database and extract time-stamped CST survey data. The data collection and integrity check process will be tested in before the EP7 is open for evaluation. In July 1996, the EP6 team will begin extracting and analyzing survey data on a weekly basis. At the end of the Evaluation Period, the data will be compiled, analyzed with data collected through usability tests, and the results will be reported in the "EP7 Results Report."

4.5 Survey Questions

Users will be asked to base their answer on the amount with which they agree or disagree with each question. A Likert scale of 1 (Strongly Disagree) to 5 (Strongly Agree) will be used to quantify users' responses.

At the time this document went to press, the survey questions were not available for publication. A sample of survey questions which were used in the evaluation of EP6 are provided below.

General

- Navigation through EP6 was easy.
- The EP6 window layouts are easy to understand.
- I like the way the EP6 Help function is implemented.
- The Help provided in EP6 was adequate for my needs.
- The Help provided in EP6 was easy to understand.
- It was easy to fill out the user survey.
- The layout of the survey was clear and easy to use.
- My experience using EP6 was positive.

This page intentionally left blank.

5. Usability Testing

5.1 Test Methodology

In addition to the use of the Comment Survey Tool by all EP7 Evaluators, a subset of the EP7 Evaluators will perform usability testing at the Landover EDF. Usability testing is a method of quantifying the ease with which a system can be used. It has been utilized to capture various measures such as user satisfaction, access time, and error rates. These are all used to locate areas of the system in need of improvement. The methodology employed for usability testing of EP7 is adapted from a paper written by Martha Szczur, "Usability Testing on a Budget." Szczur describes an efficient and low cost method of testing and quantifying usability. This methodology was employed for the evaluation of EP3, EP4, and EP6.

A five step process will be used to design the usability test and carry out the test series on EP7. Before usability testing can begin the goals of the testing and method of measuring them were identified, section 4.2. Following this, task areas were analyzed and specific user tasks clarified, section 4.6. Third, the sample test participant pool was defined and targeted, section 4.3. Once the usability test packet is completed it will be tested as part of a validation process to locate any areas where the test does not flow logically, is unclear, or is unduly lengthy. This step helps to determine where the usability test participants may become confused, lost, or frustrated. The final step will be to administer the test to the usability test participants.

5.2 Test Goals

EP7 will be released for user evaluation on July 30 and will continue through September 15, 1996. During this Evaluation Period, users will be able to travel to the Landover ECS Development Facility to evaluate both the X/Motif version and Web-based version of EP7.

The usability testing goals for EP7 include:

1. Quantify relative ease of use of tested EP7 functions.
2. Identify whether or not tested EP7 functions are intuitive to perform.
3. Identify, quantify, and prioritize functions or features of EP7 that could benefit from improvement. This goal includes:
 - evaluate the Human-computer interface (HCI) designs.
 - evaluate the data search, identification, and selection features.
 - evaluate the data organization features.
 - evaluate the data viewing features.
 - evaluate the HTML/WWW related features

- evaluate the user customization features
- 4. Provide alternative solutions to improve these features.
- 5. Provide feedback to developers.

5.3 Test Participants

A number of researchers and scientists have volunteered to become usability test participants; these are a subset of the larger EP7 Evaluator group. Participants will be divided into two user classes depending upon their level of expertise: Novice and Expert. The Expert class will consist of those who are familiar with the system and were involved in the design and construction of PW2, EP6, and EP7. The Novice class will contain those who are unfamiliar with PW2, EP6, and EP7 or who have used them on a very limited basis. In order to determine the category into which each Participant falls, they will be asked to answer a few questions regarding their computer experience. Participants will be guaranteed anonymity and encouraged to comment on any and all aspects of the system, its design, and implementation.

In addition to the test Participants, other personnel are involved in each usability test session. Each test session will include at least two Observers. Their responsibilities include taking notes of Participant reaction, comments, and body language. A Facilitator will conduct the usability test, act as host to the test Participants, and keep time for the test session. The Facilitator will keep notes and coordinate with the Observers to run a smooth and efficient test session. The records kept by the Observers and the Facilitator will be combined with task times, and user survey results for analysis.

5.4 Test Schedule

Each usability test session will be allocated a minimum of two hours. This will give the EP7 Facilitator and Observers time to set up for the session and introduce the Participants to the Observers, to the usability testing environment, and inform them of the test schedule. After introductions are completed, Participants will be given a short demonstration of EP7 showing each of the capabilities, and how to access them. Participants will be given a chance to ask questions before beginning the usability test.

After the usability tasks are completed the users will take the on-line survey using the CST and be able to register their comments in the free-text field provided at the bottom of each CST window. At the conclusion of the test, Participants will be presented with their choice of a candy treat as a thank you for participating in the usability test. Once the Participants have left, the Facilitator and Observers will debrief and then transcribe their notes from the session.

5.5 Test Environment

Usability tests should be conducted in a quiet area free from noise, movement of people and goods, and other distractions. It is preferred that tests be conducted in a room allocated for that purpose, or in an area of a larger room that has been restricted to those participating in the EP7 Evaluation. There should be room enough for two test Participants seated side by side at the

computer, a Facilitator sitting next to them, and two Observers located behind the test Participants. The environment should allow for sufficient Participant elbow room and an area for the Facilitator note taking. See figure below for test environment.

Although usability tests are usually conducted using one Participant at a time, it is not uncommon for two Participants to evaluate a system simultaneously. Usability test sessions for EP4 and EP6 often included Participants working in pairs, this was found to be very successful. This arrangement allows for increased interaction between the Participants and an increase in the number and detail of comments generated by the Participants. Another advantage of conducting tests using Participant pairs is that they can help each other navigate the system and learn how to use various features, rather than turning to the Facilitator each time they become "stuck." This will help to keep the usability test running smoothly and maintain Participant interest.

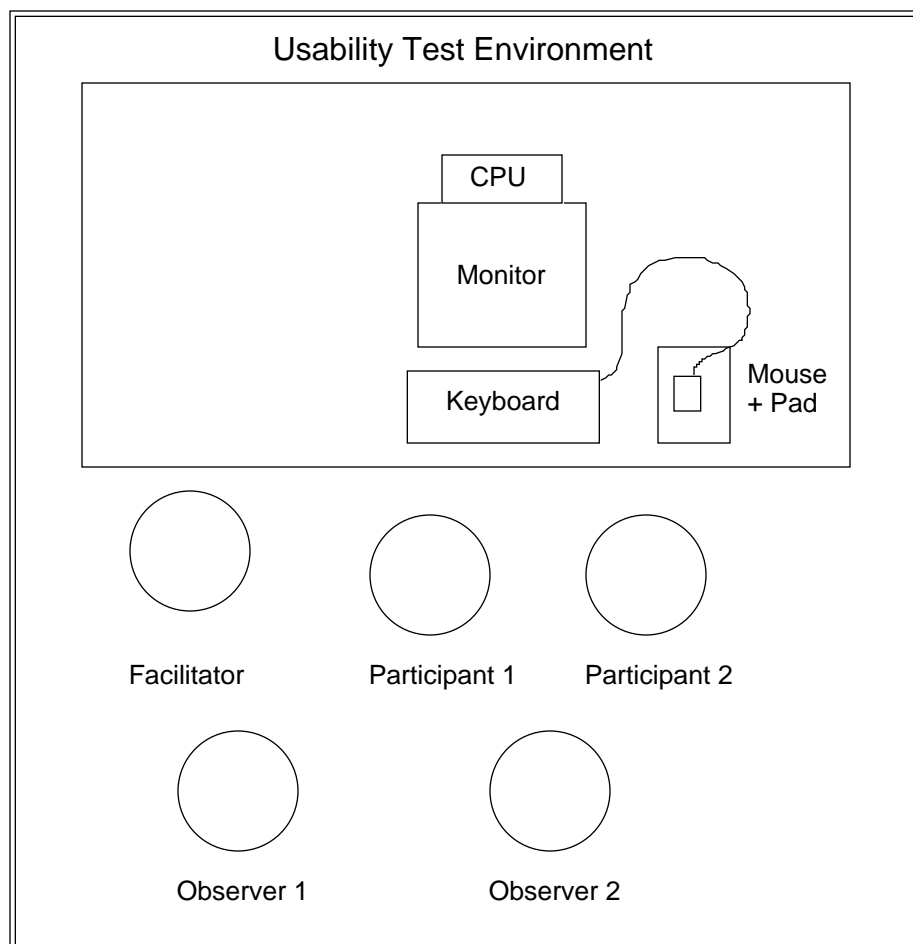


Figure 5-1. Usability Test Environment

5.6 Test Tasks

The EP6 developers and others involved with EP7 development provided a suggested list of tasks that they would like the users to execute. These tasks will be incorporated into a usability test packet draft and reviewed, as were the CST survey questions, to determine if they and the test packet are suitable. By suitable, the tasks:

- should flow from one task to the next in logical order.
- should allow the users to access and evaluate the significant portions of EP7, especially those portions of EP7 that developers highlighted specifically for evaluation.
- should be combined in a task packet so as to vary the difficulty from one task to the next. This will help to keep user interest focused on the test and reduce participant frustration.
- once combined in a test packet, should not take more than one hour to complete.
- should complement and coordinate with the CST survey questions.

5.7 Coordination with Comment Survey Tool

Data collected from Usability Participants during tests and their responses in the Comment Survey Tool will be compared with the data provided by the Remote EP7 Evaluators gathered in the Comment Survey Tool. Comparing the results of the usability tests with the results gathered from the Remote Evaluators will allow for in-depth analysis of user preferences, user satisfaction, user needs and whether they were met by EP7. It is expected that both Usability Participants and Remote Evaluators will have similar experiences and comments about the EP7. However, it is likely that the Remote Evaluators will have more comments regarding EP7 Help documentation, access, and system performance, than the Participants in usability testing. In addition, the Remote Evaluators will be able to log on to EP7 and leave their comments in the CST a number of times during the course of the evaluation period. It will be interesting to monitor how, or if their comments change as a result of becoming more familiar with the EP7.

In order to make the coordinated data analysis more efficient a data storage and analysis plan will be devised prior to the Evaluation Period. For example, information such as statistics to be collected, graphics to be produced, level of detail to attain, etc. will be identified, this will save time and energy during the data collection process and also during data analysis. The data analysis of the usability testing and the Remote Evaluators results will each be summarized in the EP7 Results Report, to be published in October 1996.

6. Demonstration Scripts

6.1 Overview Demonstration for Usability Participants

Prior to conducting each usability test the Participants will be shown a brief demonstration of EP7, tailored to the version (X/Motif or Web) they will evaluate. The Overview Demonstration script will contain basic information for the Participants to learn how to access and operate EP7 functions and locate resources. However, the script will not go in to great detail because that would bias the Participants' response and actions in the usability test. The Overview Demonstration for the X/Motif portion of EP7 will show Participants files on the EP7 Desktop/Workbench being manipulated, how to access and select search criteria in the ESST, how to access the Data Dictionary, download information from the Advertising Service, lookup a document in the Document Search Tool, and access the Comment Survey Tool.

An outline of the Overview Demonstration for X/Motif is listed below:

- 1) Login to the EP7 and view the Desktop/Workbench
- 2) Describe Workbench, explain how to navigate through it, show the help, pull-down menus, etc.
- 3) Access the ESST and show features of the tool
- 4) Access the Data Dictionary to look up the definition of an Earth Science term
- 5) Search for information in the Advertising Service and install the resulting service/provider it on the Desktop/Workbench
- 6) Search for a document in the Document Search Tool
- 7) Access the Comment Survey Tool
- 8) Return to the Desktop/Workbench

6.2 Remote Evaluator Demonstration

Remote EP7 Evaluators will also have access to the demonstration script which will be posted on the EP7 Home Page. Because the EP7 Evaluators will not have direct access to an EP7 developer or a usability test Facilitator, a detailed EP7 demonstration script will be provided for them. Just as Usability Participants will be given a context, or scenario within which to evaluate EP7, the Remote Evaluator Demonstration script will be written with a similar context or scenario in mind so that Remote Evaluators may have a context within which to evaluate the software.

The Remote Evaluator Demonstration script will contain information on where various features are located in EP7, how to access them, and some of their more detailed capabilities. For the

most part it will follow the outline of the Overview Demonstration for Usability Participants script but in much more detail. The Overview script will be provided in "actions and associated results" format, with occasional material added to further clarify actions/results. The Evaluator Demonstration script is designed for individuals to follow on their own.

7. References

Day (1996), Evaluation Package 6 (EP6) Results Report, Technical Paper, 160-TP-010-001.

HAIS, (1994), EP3 Evaluation Report, Technical Paper, 194-TP-441-001 (formerly document number, 194-0441).

Nielsen, Jakob (1994), Usability Engineering, User Interface Strategies '94, University of Maryland at College Park.

Poston (1994), EP4 Evaluation Report, Technical Paper, 222-TP-004-001

Szczur, Martha (1993), Usability Testing - On a Budget: a usability test case study, NASA/GSFC.

This page intentionally left blank.